Biedermann, Stefanie; Dette, Holger; Hoffmann, Philipp

Working Paper

Constrained optimal discriminating designs for Fourier regression models

Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen, No. 2006,28

Provided in cooperation with:
Technische Universität Dortmund

Constrained optimal discriminating designs for Fourier regression models

Stefanie Biedermann, Holger Dette, Philipp Hoffmann
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

June 10, 2006

Abstract

In this article, the problem of constructing efficient discriminating designs in a Fourier regression model is considered. We propose designs which maximize the efficiency for the estimation of the coefficient corresponding to the highest frequency subject to the constraints that the coefficients of the lower frequencies are estimated with at least some given efficiency. A complete solution is presented using the theory of canonical moments, and for the special case of equal constraints the optimal designs can be found analytically.

AMS Subject classification: 62K05, 62J05
Keywords and Phrases: Constrained optimal designs, trigonometric regression, D_1-optimal designs, Chebyshev polynomials, canonical moments.

1 Introduction

The Fourier regression or trigonometric regression model

$$g_{2d}(x) = a_0 + \sum_{j=1}^{d} a_j \sin(jx) + \sum_{j=1}^{d} b_j \cos(jx) \quad x \in [-\pi, \pi] \tag{1.1}$$

$$g_{2d-1}(x) = a_0 + \sum_{j=1}^{d} a_j \sin(jx) + \sum_{j=1}^{d-1} b_j \cos(jx) \quad x \in [-\pi, \pi] \tag{1.2}$$

is widely used to describe periodic phenomena [see, e.g., Mardia (1972), Kitsos, Titterington and Torsney (1988) or the recent collection of research papers in biology edited by Lestrel (1997)]. The value $2d$ in (1.1) or $2d-1$ (1.2) is usually denoted as the degree of the Fourier regression model and $a_0, a_1, \ldots, a_d, b_1, \ldots, b_d$ denote unknown parameters, which have to be estimated from the data. The problem of designing experiments for models of the form (1.1) has been discussed by several authors; see, e.g., Karlin and Studden (1966), page 347, Fedorov (1972), page 94, Hill (1978), Lau and Studden (1985) for optimal designs on the full circle, and some recent contributions by Dette, Melas and Biedermann (2002) and Dette, Melas and Pepelyshev (2002) for optimal designs on a partial circle. Most authors concentrate on the problem of determining optimal designs for the estimation of the full vector of unknown parameters, whereas the problem of constructing optimal designs for model discriminating has only recently been considered in the literature [see Dette and Haller (1998), Dette and Melas (2003) and Zen and Tsai (2004)]. The present paper is devoted to the problem of constructing optimal discriminating designs using constrained optimality criteria.

Constrained optimal designs have primarily been considered by Stigler (1971), Studden (1982b) and Lee (1988a, b), whereas the more recent work by Cook and Wong (1994), Dette (1995) and Clyde and Chaloner (1996) investigates the relation between this approach and compound optimality criteria. Although these results are interesting from a theoretical point of view constrained optimal designs have to be found numerically and explicit results can only be inferred in rare cases. Recently, Dette and Franke (2000) characterized constrained optimal discriminating designs for polynomial regression models utilizing the theory of canonical moments, which was introduced by Skibinsky (1967) and applied by Studden (1980, 1982a, 1982b, 1989) for determining optimal designs in polynomial regression models. The problem of finding constrained optimal discriminating designs for Fourier regression models, however, has not been considered yet.

The present paper is devoted to this problem. For the construction of constrained optimal designs we assume that the highest frequency of the model has been fixed and determine the design such that the coefficient corresponding to this frequency is estimated with maximal efficiency subject to the constraints that the coefficients corresponding to the highest frequencies in the models of lower degree can be estimated with some guaranteed efficiency. The optimality criterion is carefully described in Section 2. In Section 3 we briefly review some facts from the theory of canonical moments [see Dette and Studden (1997)], which is the basic tool for the construction of optimal discriminating designs. A complete characterization of the constrained optimal descriminating designs is given, and in the special case of equal bounds the optimal designs can be found explicitly. Finally, in Section 4 we illustrate the method by several examples.

2 Constrained optimal designs in Fourier regression models

For $k = 0, \ldots, 2d$ we define

$$f_k(x) = \begin{cases} (1, \sin(x), \cos(x), \ldots, \sin(jx), \cos(jx))^T, & \text{if } k = 2j \\ (1, \sin(x), \cos(x), \ldots, \sin((j-1)x), \cos((j-1)x), \sin(jx))^T, & \text{if } k = 2j-1 \end{cases}$$

and

$$\theta_k = \begin{cases} (a_0, a_1, b_1, \dots, a_j, b_j)^T & \text{if } k = 2j \\ (a_0, a_1, b_1, \dots, a_{j-1}, b_{j-1}, a_j)^T & \text{if } k = 2j-1 \end{cases}$$

then the models in (1.1) and (1.2) can be written as $g_k(x) = f_k(x)^T \theta_k$ where $k = 2d$ or $k = 2d-1$, respectively. An approximate design is a probability measure σ with finite support on the interval $[-\pi, \pi]$ with the interpretation that observations are taken at the support points in proportion to the corresponding masses. The analogue of the matrix $X^T X$ in the Fourier regression model $g_k(x)$ is the information matrix

$$M_k(\sigma) = \int_{-\pi}^{\pi} f_k(x) f_k^T(x) d\sigma(x). \tag{2.1}$$

An optimal design maximizes an appropriate real-valued function of the information matrix [see Pukelsheim (1993)], and there are numerous criteria which can be used for the characterization of efficient designs. Our optimality criterion for constructing discriminating designs is motivated by a multiple testing procedure [see e.g. Anderson (1994)], where, starting with the given regression $g_{2d}(x)$ or $g_{2d-1}(x)$ in (1.1) or(1.2), respectively, one tests the hypotheses $H_0^{(2d)} : \theta_{2d} = 0$, $H_0^{(2d-1)} : \theta_{2d-1} = 0, \dots, H_0^{(1)} : \theta_1 = 0$ successively, and decides for the model g_{k_0} where k_0 is the first index for which the hypothesis $H_0^{(k_0)} : \theta_{k_0} = 0$ is rejected. (Note that this sequence of tests can be stopped earlier if the minimal degree of the Fourier regression model is pre-specified.) The quantities corresponding to the noncentrality parameter of the F-test for the hypothesis $H_0^{(k)}$ are given by

$$\delta_k(\sigma) = (e_k^T M_k^{-1}(\sigma) e_k)^{-1} \quad k = 1, \dots, 2d, \tag{2.2}$$

where e_k denotes the $(k+1)$th unit vector in $\mathbb{R}^{k+1}$ and the design σ is assumed to have at least $(2d+1)$ support points [see Pukelsheim (1993), p. 70]. A design σ_k^* is called D_1-optimal if it maximizes δ_k, and the expression

$$\text{eff}_k(\sigma) := \frac{\delta_k(\sigma)}{\delta_k(\sigma_k^*)}, \quad k = 1, \dots, 2d \tag{2.3}$$

is called the D_1-efficiency of the design σ in the Fourier regression model $g_k(x)$. Recall that a design maximizing δ_k is optimal for discriminating between the trigonometric regression models g_{k-1} and g_k. Dette and Haller (1998) proposed to maximize a weighted p-mean of the quantities $\delta_1, \dots, \delta_{2d}$ for the construction of an optimal design for discriminating between the models $\{g_1, \dots g_{2d}\}$. In the present paper, we propose an alternative optimality criterion to obtain efficient discriminating designs. This approach is attractive if the main interest of the experimenter is in estimating the coefficient corresponding to the highest frequency, but the design should also allow an efficient discriminating between the models of lower degree.

We consider two cases for determining a constrained optimal discriminating design for the Fourier regression model. The first approach considers the highest cosine frequency as most important and a constrained optimal discriminating design σ^* maximizes

$$\text{eff}_{2d}(\sigma) \quad \text{subject to} \quad \text{eff}_l(\sigma) \geq c_l, \quad l = 2d-1, 2d-2, \dots, 2d-2j-1 \tag{2.4}$$

for some $j \in \{0, \ldots, d-1\}$, whereas the second criterion determines the design which maximizes

$$\text{(2.5)} \quad \text{eff}_{2d-1}(\sigma) \quad \text{subject to the constraints} \quad \text{eff}_l(\sigma) \geq c_l, \quad l = 2d, 2d-2, \ldots, 2d-2j-1$$

where $j \in \{1, \ldots, d-1\}$, and the quantities $c_{2d-2j-1}, \ldots, c_{2d} \in (0,1)$ are given by the experimenter and reflect the desired precision for the estimation of the coefficient corresponding to the highest frequency in the models $g_{2d-2j-1}, \ldots, g_{2d}$. For the solution of this constrained optimization problem we need several tools, which will be explained in what follows.

It follows by standard arguments [see Pukelsheim (1993), Chap. 4, 5] that eff_l is a concave function on the set of designs on the interval $[-\pi, \pi]$ and invariant with respect to a reflection of the design σ at the origin. Consequently, if there exists an optimal constrained optimal discriminating design, then there also exists an optimal design in the set Σ of all symmetric designs on the interval $[-\pi, \pi]$. We note that these symmetric designs induce designs ξ_σ on the interval $[-1, 1]$ by the projection

$$\xi_\sigma(\cos x) = \begin{cases} 2\sigma(x) = 2\sigma(-x) & \text{if} \quad 0 < x \leq \pi \\ \sigma(0) & \text{if} \quad x = 0 \end{cases} \tag{2.6}$$

for any symmetric design $\sigma \in \Sigma$. The corresponding set of the measures ξ_σ on $[-1, 1]$ will be denoted by $\Sigma_{[-1,1]}$. It was shown in Dette and Haller (1998) that for any $\sigma \in \Sigma$

$$\delta_k(\sigma) = \begin{cases} 2^{2(j-1)} \frac{|A_j(\xi_\sigma)|}{|A_{j-1}(\xi_\sigma)|} & \text{if} \quad k = 2j \\ 2^{2(j-1)} \frac{|B_j(\xi_\sigma)|}{|B_{j-1}(\xi_\sigma)|} & \text{if} \quad k = 2j-1. \end{cases} \tag{2.7}$$

where $B_0(\xi_\sigma) = A_0(\xi_\sigma) = 1$, and

$$A_k(\xi_\sigma) = \Big(\int_{-1}^{1} z^{i+j} d\xi_\sigma(z) \Big)_{i,j=0}^{k} \tag{2.8}$$

$$B_k(\xi_\sigma) = \Big(\int_{-1}^{1} (1-z^2) z^{i+j} d\xi_\sigma(z) \Big)_{i,j=0}^{k-1} \tag{2.9}$$

denote the information matrices of the design ξ_σ on the interval $[-1, 1]$ for a homoscedastic and a heteroscedastic polynomial regression model with efficiency function $\lambda(z) = (1 - z^2)$ [see Karlin and Studden (1966)], respectively. Consequently, the problem of determining constrained optimal discriminating designs for the Fourier regression model can be solved by maximizing a certain function over the set of probability measures on the interval $[-1, 1]$ and transforming the maximizing measure back via (2.6).

The problem of maximizing the right hand side of (2.7) over the set $\Sigma_{[-1,1]}$ if $k = 2j$ is in fact the D_1-optimal design for the ordinary polynomial regression model, while for odd values of k the right hand side of (2.7) corresponds to the weighted polynomial regression with efficiency function $\sigma^2(x) = \sigma^2/(1-x^2), x \in (-1, 1)$. The solutions of these problems are well known [see Studden (1968, 1982b)] and yield $\delta_k(\sigma_k^*) = \max_\sigma \delta_k(\sigma) = 1$ $(k = 1, \ldots, 2d)$, and therefore the efficiency of a symmetric design σ defined in (2.3) can be rewritten as

$$\text{eff}_k(\sigma) = \begin{cases} 2^{2(j-1)} \frac{|A_j(\xi_\sigma)|}{|A_{j-1}(\xi_\sigma)|} & \text{if } k = 2j \\ 2^{2(j-1)} \frac{|B_j(\xi_\sigma)|}{|B_{j-1}(\xi_\sigma)|} & \text{if } k = 2j-1. \end{cases} \tag{2.10}$$

3 The solution of the constrained optimal design problem

For the characterization of the measure $\xi_{\sigma^*} \in \Sigma_{[-1,1]}$ corresponding to the constrained optimal discriminating design σ^* by the relation (2.6) we require some basic facts about the theory of canonical moments which has been introduced by Studden (1980, 1982a,b) in the context of optimal design. We will only give a very brief heuristical introduction and refer to the monograph of Dette and Studden (1997) for more details.

It is well known that a probability measure on the interval $[-1,1]$, say ξ, is determined by its sequence of moments $(m_1, m_2, \ldots)$. Skibinsky (1967) defined a one to one mapping from the sequences of ordinary moments onto sequences $(p_1, p_2, \ldots)$ whose elements vary independently in the interval $[0,1]$. For a given probability measure on the interval $[-1,1]$ the element p_j of the corresponding sequence is called the jth canonical moment of ξ. If j is the first index for which $p_j \in \{0,1\}$ then the sequence of canonical moments terminates at p_j, and the measure is supported at a finite number of points. The support points and corresponding masses can be found explicitly by evaluating certain orthogonal polynomials [see Dette and Studden (1997), Chapter 3]. The set of probability measures on the interval $[-1,1]$ with first k canonical moments equal to $(p_1, \ldots, p_k) \in (0,1)^{k-1} \times [0,1]$ is a singleton if and only if $p_k \in \{0,1\}$. Otherwise there exists an uncountable number of probability measures corresponding to $(p_1, \ldots, p_k)$ [see Skibinsky (1986)].

It turns out that the determinants in (2.10) can be described in terms of the canonical moments $p_1, p_2, \ldots$ of the measure ξ_σ [see Studden (1982b)], that is

$$|A_k(\xi_\sigma)| \;=\; 2^{k(k+1)} \prod_{\ell=1}^{k} (q_{2\ell-2}p_{2\ell-1}q_{2\ell-1}p_{2\ell})^{k-\ell+1} \tag{3.1}$$

$$|B_k(\xi_\sigma)| \;=\; 2^{k(k+1)} \prod_{\ell=1}^{k} (p_{2\ell-2}q_{2\ell-1}p_{2\ell-1}q_{2\ell})^{k-\ell+1} \tag{3.2}$$

where $p_0 = 1,\; q_0 = 1$ and $q_j = 1 - p_j$ for $j \geq 1$. Observing (2.10), (3.1) and (3.2), we find that the efficiencies are increasing functions of the terms $p_{2j-1}q_{2j-1}$, and consequently the canonical moments of the projection ξ_{σ^*} of the constrained optimal discriminating design σ^* satisfy

$$p_{2\ell-1} \;=\; \frac{1}{2} \qquad \ell = 1, \ldots, d. \tag{3.3}$$

Therefore we can restrict ourselves to designs with this property, and (2.10) reduces to

$$\mathrm{eff}_k(\sigma) = \begin{cases} 2^{2j-2} p_{2j} \prod_{\ell=1}^{j-1} q_{2\ell}p_{2\ell} & \text{if } k = 2j \\ 2^{2j-2} q_{2j} \prod_{\ell=1}^{j-1} q_{2\ell}p_{2\ell} & \text{if } k = 2j-1 \end{cases} \tag{3.4}$$

where $p_2, p_4, \ldots$ denote the canonical moments of even order of the design $\xi_\sigma \in \sum_{[-1,1]}$ satisfying (3.3) and corresponding to the measure σ via (2.6). Our main result gives a characterization of the canonical moments of ξ_{σ^*}.

Theorem 3.1.
(a) *If there exists a constrained optimal discriminating design for the vector* $c_{2d-2j-1}, \ldots, c_{2d-1}$

in (2.4) , then there also exists a symmetric optimal discriminating design σ^*. The canonical moments up to the order $2d$ of the corresponding projection ξ_{σ^*} are determined by the system of equations

$$
\begin{aligned}
p_{2n-1} &= \frac{1}{2} \qquad n = 1, \ldots, d \\
p_{2n} &= \frac{1}{2} \qquad n = 1, \ldots, d-j-1 \\
p_{2d-2j+2n} &= \begin{cases} 1 - \max\{\frac{1}{2}, \dfrac{c_{2d-2j+2n-1}}{2^{2n}\prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}}\}, & \text{if } c_{2d-2j+2n-1} > c_{2d-2j+2n} \\ \max\{\frac{1}{2}, \dfrac{c_{2d-2j+2n}}{2^{2n}\prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}}\}, & \text{if } c_{2d-2j+2n} \geq c_{2d-2j+2n-1} \end{cases} \\
& \qquad\qquad\qquad\qquad\qquad\qquad n = 0, \ldots, j-1 \\
p_{2d} &= 1 - \frac{c_{2d-1}}{2^{2j}\prod_{l=d-j}^{d-1} p_{2l}q_{2l}}.
\end{aligned}
$$

(b) *If there exists a constrained optimal discriminating design for the vector $c_{2d-2j-1}, \ldots, c_{2d-2}, c_{2d}$ in (2.5), then there also exists a symmetric constrained optimal discriminating design σ^*. The canonical moments up to the order $2d$ of the corresponding projection ξ_{σ^*} are determined by the system of equations*

$$
\begin{aligned}
p_{2n-1} &= \frac{1}{2} \qquad n = 1, \ldots, d \\
p_{2n} &= \frac{1}{2} \qquad n = 1, \ldots, d-j-1 \\
p_{2d-2j+2n} &= \begin{cases} 1 - \max\{\frac{1}{2}, \dfrac{c_{2d-2j+2n-1}}{2^{2n}\prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}}\}, & \text{if } c_{2d-2j+2n-1} > c_{2d-2j+2n} \\ \max\{\frac{1}{2}, \dfrac{c_{2d-2j+2n}}{2^{2n}\prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}}\}, & \text{if } c_{2d-2j+2n} \geq c_{2d-2j+2n-1} \end{cases} \\
& \qquad\qquad\qquad\qquad\qquad\qquad n = 0, \ldots, j-1 \\
p_{2d} &= \frac{c_{2d}}{2^{2j}\prod_{l=d-j}^{d-1} p_{2l}q_{2l}}.
\end{aligned}
$$

Proof. Because both parts are proved similarly, we restrict ourselves to a proof of part (a). By the previous discussion the canonical moments of odd order $1, 3, \ldots 2d-1$ must be $1/2$. Note that

$$\mathrm{eff}_{2d-2j+2n}(\sigma) = p_{2d-2j+2n} 2^{2(d-j+n-1)} \prod_{l=1}^{d-j+n-1} p_{2l}q_{2l}\ , \quad n = 0 \ldots, j.$$

In order to maximize these efficiencies we have to choose the canonical moments such that the products $p_{2l}q_{2l}$ are as large as possible. This can be accomplished by choosing p_{2l} as close as possible to the value $1/2$ such that the constraints in (2.4) are satisfied. Since there are no restrictions on the efficiencies $\mathrm{eff}_1(\sigma), \ldots, \mathrm{eff}_{2d-2j-2}(\sigma)$ we obtain $p_2 = \ldots = p_{2d-2j-2} = \frac{1}{2}$. Plugging this choice into the formulae for the higher order efficiencies, (3.4) reduces to

$$\mathrm{eff}_{2d-2j+2n-1}(\sigma) = q_{2d-2j+2n} 2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}$$

$$\mathrm{eff}_{2d-2j+2n}(\sigma) = p_{2d-2j+2n} 2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}.$$

We start with the case $n = 0$, for which the representations $\mathrm{eff}_{2d-2j-1}(\sigma) = q_{2d-2j}$, $\mathrm{eff}_{2d-2j}(\sigma) = p_{2d-2j}$ yield the constraints

$$p_{2d-2j} \geq c_{2d-2j}, \quad q_{2d-2j} \geq c_{2d-2j-1}.$$

Consequently any design ξ_σ for which $p_{2d-2j} \in [c_{2d-2j}, 1 - c_{2d-2j-1}]$ satisfies the constraints of order $2d - 2j$ und $2d - 2j - 1$. We therefore assume that $c_{2d-2j} + c_{2d-2j-1} \leq 1$ in what follows to ensure the existence of such a design. If $\frac{1}{2} \in [c_{2d-2j}, 1 - c_{2d-2j-1}]$ one can choose $p_{2d-2j} = \frac{1}{2}$ to maximize $p_{2d-2j} q_{2d-2j}$. Else we have either $c_{2d-2j} \geq \frac{1}{2}$ or $1 - c_{2d-2j-1} \leq \frac{1}{2}$, and we choose $p_{2d-2j} = c_{2d-2j}$ or $p_{2d-2j} = 1 - c_{2d-2j-1}$, respectively.
If $n > 0$ we note that the constraints $\mathrm{eff}_{2d-2j+2n}(\sigma) \geq c_{2d-2j+2n}$ and $\mathrm{eff}_{2d-2j+2n-1}(\sigma) \geq c_{2d-2j+2n-1}$ reduce to

$$p_{2d-2j+2n} \geq \frac{c_{2d-2j+2n}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}} =: c'_{2d-2j+2n}$$

$$q_{2d-2j+2n} \geq \frac{c_{2d-2j+2n-1}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}} =: c'_{2d-2j+2n-1}$$

Therefore the same arguments as presented for the case $n = 0$ yield the corresponding result for $p_{2d-2j+2n}$. Finally we consider the case $n = j$, where there is only one constraint $\mathrm{eff}_{2d-1}(\sigma) \geq c_{2d-1}$, which can be rewritten as

$$p_{2d} \leq 1 - \frac{c_{2d-1}}{2^{2d-2} \prod_{l=1}^{d-1} p_{2l} q_{2l}}.$$

In order to maximize this expression one has to choose p_{2d} such that there is equality. This proves the final assertion of part (a) in Theorem 3.1. $\square$

Remark 3.2. Note that Theorem 3.1 characterizes the canonical moments up to the order $2d$ of the projection ξ_{σ^*} of the (symmetric) constrained optimal discriminating design σ^*. In general $p_{2d} \notin \{0, 1\}$ and in these cases there exists an infinite number of probability measures on the interval $[-1, 1]$ with the canonical moments $p_1, \ldots p_{2d}$ [see Skibinsky (1986)]. Each of these measures corresponds to a constrained optimal discriminating design by the projection (2.6).
Before we illustrate this phenomenon in the following section we present two further results, where the solution of the constrained optimal design problem can be found explicitly. For this purpose let $T_j(x)$ and $U_j(x)$ denote the jth Chebyshev polynomial of the first and second kind, respectively [see Rivlin (1974)].

Theorem 3.3. *Consider the constrained optimal design problem in (2.4) where* $c_{2d-2j} = \ldots = c_{2d-2} = c \in (0,1)$*,* $c_l < c$ $(l = 2d-2j-1, \ldots 2d-1)$*. If there exists a constrained optimal discriminating design, then there also exists a symmetric constrained optimal discriminating design* σ^**.*
(a) *If* $c > 1/2$*, define*

$$\kappa = \kappa(c_{2d-1}, c) = \frac{1}{2} - \frac{c_{2d-1}}{4c} + \frac{2c-1}{2((2c-1)j-2c)},$$

$$P_{d+1}^*(x) = (xU_j(x) - 2\kappa U_{j-1}(x))T_{d-j}(x) - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))T_{d-j-1}(x), \tag{3.5}$$

$$P_d^{(1)}(x) = (xU_j(x) - 2\kappa U_{j-1}(x))U_{d-j-1}(x) - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))U_{d-j-2}(x) \tag{3.6}$$

The polynomial $P_{d+1}^*(x)$ *has* $d+1$ *distinct roots* $x_0, \ldots x_d$ *in the interval* $(-1,1)$*, and the design* ξ_{σ^*} *with masses*

$$\lambda_k = \frac{P_d^{(1)}(x_k)}{\frac{d}{dx}P_{d+1}^*(x)|_{x=x_k}}, \quad k = 0\ldots,d \tag{3.7}$$

at $x_0, \ldots, x_d$ *corresponds to a constrained optimal discriminating design for the optimization problem (2.4) by the projection (2.6).*
(b) *If* $c < 1/2$*, define*

$$P_{d+1}^*(x) = T_{d+1}(x) - (1-2c)T_{d-1}(x),$$
$$P_d^{(1)}(x) = U_d(x) - (1-2c)U_{d-2}(x).$$

The polynomial $P_{d+1}^*(x)$ *has* $d+1$ *distinct roots* $x_0, \ldots x_d$ *in the interval* $(-1,1)$*, and the design* ξ_{σ^*} *with masses (3.7) at* $x_0, \ldots, x_d$ *corresponds to a constrained optimal discriminating design for the optimization problem (2.4) by the projection (2.6).*

Remark 3.4. It will be shown in the proof of Theorem 3.3 that a necessary condition for the existence of a symmetric constrained optimal discriminating design for the design problem (2.4) is $\kappa > 0$, which ensures that the value of the canonical moment p_{2d} will be within the interval $(0,1)$. If $\kappa \leq 0$ it is therefore recommended to modify the choices of c and c_{2d-1} accordingly so that κ attains a positive value, before starting to calculate the optimal design.

Proof of Theorem 3.3. We only prove part (a) of the Theorem. Part (b) follows by similar (and even simpler) arguments. Note that the canonical moments of the constrained optimal discriminating design can be obtained by Theorem 3.1. The canonical moments of odd order satisfy

$$p_{2n-1} = \frac{1}{2}, \ n = 1, \ldots d, \tag{3.8}$$

while the canonical moments of even order less or equal than $2d-2j-2$ are given by

$$p_{2n} = \frac{1}{2}, \ n = 1, \ldots d-j-1. \tag{3.9}$$

For the next canonical moment of even order we have from Theorem 3.1

$$p_{2d-2j} = \max\{\frac{1}{2}, c\} = c,$$

and it can be shown by a straightforward induction that

$$p_{2d-2j+2t} = \frac{1}{2}\frac{(2c-1)t-2c}{(2c-1)(t+1)-2c}\ , \quad t = 0,\ldots,j-1. \tag{3.10}$$

Note that this representation implies

$$\frac{1}{2} < c < \frac{j+1}{2j},$$

because the canonical moments vary in the interval $(0,1)$. The remaining canonical moment of order $2d$ is obtained by a direct calculation, that is

$$p_{2d} = 1 - \frac{1}{2}\frac{c_{2d-1}}{c}\frac{(2c-1)j-2c}{(2c-1)(j+1)-2c}. \tag{3.11}$$

It follows from a straightforward but tedious calculation that $p_{2d} \in (0,1) \Leftrightarrow \kappa > 0$, which proves Remark 3.4. Note that (3.8) - (3.11) do not determine a design on the interval $[-1,1]$ (except in the case $c_{2d-1} = 0$, which is excluded). In order to obtain a design with finite support we extend this sequence by

$$p_{2d+1} = \frac{1}{2}\ , \quad p_{2d+2} = 0, \tag{3.12}$$

and note that the design ξ^*_σ on the interval $[-1,1]$ with canonical moments (3.8) - (3.12) is uniquely determined and has $d+1$ support points [see Skibinsky (1986)]. For the calculation of the support points and corresponding weights we apply Theorem 3.6.1 in Dette and Studden (1997). By this result the design ξ_{σ^*} has weights

$$\lambda_k \ = \ \frac{P_d^{(1)}(x_k)}{\frac{d}{dx}P^*_{d+1}(x)|_{x=x_k}} \tag{3.13}$$

at the roots $x_0, \ldots x_d$ of the polynomial $P^*_{d+1}(x)$, where $P^*_{d+1}(x)$ and $P_d^{(1)}(x)$ are obtained from the recursion

$$W_{k+1}(x) \ = \ xW_k(x) - q_{2k-2}p_{2k}W_{k-1}(x) \tag{3.14}$$

(note that $p_{2j-1} = \frac{1}{2}$ for $j = 1,\ldots,d+1$) with different initial conditions, that is

$$P^*_{d+1}(x) \ = \ W_{d+1}(x) \quad \text{for} \quad W_{-1}(x) \equiv 0, \qquad W_0(x) \equiv 1 \tag{3.15}$$

$$P_d^{(1)}(x) \ = \ W_{d+1}(x) \quad \text{for} \quad W_0(x) \equiv 0, \qquad W_1(x) \equiv 1. \tag{3.16}$$

We now calculate these polynomials using (3.9) - (3.12) and begin with $P^*_{d+1}(x)$. ¿From the initial condition in (3.15) and (3.9) we obtain by a straightforward calculation

$$W_{d-j}(x) = \frac{1}{2^{d-j-1}}T_{d-j}(x)\ , \quad W_{d-j-1}(x) = \frac{1}{2^{d-j-2}}T_{d-j-1}(x) \tag{3.17}$$

Observing (3.10) and

$$\begin{aligned} q_{2l-2}p_{2l} &= \left(1-\frac{1}{2}\frac{(2c-1)(l-1-d+j)-2c}{(2c-1)(l-d+j)-2c}\right)\frac{1}{2}\frac{(2c-1)(l-d+j)-2c}{(2c-1)(l-d+j+1)-2c} \\ &= \frac{1}{4}\frac{(2c-1)(l-d+j+1)-2c}{(2c-1)(l-d+j+1)-2c} = \frac{1}{4} \end{aligned}$$

$(d-j<l\le d-1)$, we obtain the recursion

$$\begin{aligned} W_{d-j+1} &= xW_{d-j}(x) - \frac{1}{2}cW_{d-j-1}(x) \\ W_{l+1}(x) &= xW_l(x) - \frac{1}{4}W_{l-1}(x) \quad \text{if } d-j<l\le d-1. \end{aligned}$$

Now a straightforward induction yields

$$W_{d-j+l}(x) = \frac{1}{2^{l+d-j-1}}\left(U_l(x)T_{d-j}(x) - 2cU_{l-1}(x)T_{d-j-1}(x)\right), \quad l=1,\dots j.$$

We finally note that by (3.10) and (3.11) we have $q_{2d-2}p_{2d}=\kappa$, from which it follows that

$$P^*_{d+1}(x) = \frac{1}{2^{d-1}}\Big[(xU_j(x)-2\kappa U_{j-1}(x))T_{d-j}(x) - 2c(xU_{j-1}(x)-2\kappa U_{j-2}(x))T_{d-j-1}(x)\Big]$$

using (3.14) and (3.15). Observing the initial conditions in (3.16) it follows that the polynomial $P^{(1)}_d(x)$ can be calculated analogously, where (3.17) is replaced by

$$W_{d-j}(x) = \frac{1}{2^{d-j-1}}U_{d-j-1}(x), \quad W_{d-j-1}(x) = \frac{1}{2^{d-j-2}}U_{d-j-2}(x).$$

Consequently, by a straightforward induction we obtain

$$P^{(1)}_d(x) = \frac{1}{2^{d-1}}\Big[(xU_j(x)-2\kappa U_{j-1}(x))U_{d-j-1}(x) - 2c(xU_{j-1}(x)-2\kappa U_{j-2}(x))U_{d-j-2}(x)\Big],$$

and the assertion (a) of Theorem 3.1 follows from Theorem 3.6.1 in Dette and Studden (1997). □

We conclude this section with an analogue for the optimization problem (2.5). The proof is similar and omitted for brevity.

Theorem 3.5. *Consider the constrained optimal design problem in (2.5) where $c_{2d-2j}=\dots=c_{2d-2}=c_{2d}=c\in(0,1)$, $c_l<c$ $(l=2d-2j-1,\dots 2d-3)$. If there exists a constrained optimal discriminating design then there also exists a symmetric constrained optimal descriminating design σ^**
(a) *If $c>1/2$, define*

$$\kappa = \kappa(c_{2d},c) = \frac{c_{2d}}{4c},$$

and consider for this κ the polynomials $P^*_{d+1}(x)$ and $P^{(1)}_d(x)$ defined by (3.5) and (3.6), respectively. The polynomial $P^*_{d+1}(x)$ has $d+1$ distinct roots $x_0, \dots, x_d$ in the interval $(-1,1)$, and the design ξ_{σ^*} which has masses (3.7) at the points $x_0, \dots x_d$ corresponds to a constrained optimal discriminating design for the optimization problem (2.5) by the projection (2.6).
(b) *If $c < 1/2$, define*

$$\begin{aligned} P^*_{d+1}(x) &= T_{d+1}(x) + (1-2c)T_{d-1}(x), \\ P^{(1)}_d(x) &= U_d(x) + (1-2c)U_{d-2}(x), \end{aligned}$$

The polynomial $P^*_{d+1}(x)$ has $d+1$ distinct roots $x_0, \dots x_d$ in the interval $(-1,1)$, and the design ξ_{σ^*} with masses (3.7) corresponds to a constrained optimal discriminating design for the optimization problem (2.5) by the projection (2.6).

4 Examples

To fix ideas we will give several examples for scenarios of practical relevance. Since trigonometric regression models with degree greater than $d = 3$ are rarely considered in practice we restrict ourselves to several cases where $d = 2$ and $d = 3$. Furthermore, we present only examples for the design problem (2.4). The calculations for the problem (2.5) are similar, and examples for this problem are therefore omitted for the sake of brevity.

Example 4.1. Let us first consider the situation where $d = 2$ and $j = 1$, and the goal is to maximize the efficiency $\text{eff}_4(\sigma)$ for estimating the coefficient of the highest cosine frequency $\cos(2x)$ in model $g_4(x)$. Let now the pre-specified minimal efficiencies for estimating the highest coefficients in the lower order models be given by $c_2 = c = 0.6$ and $c_3 = 0.5$.
¿From Theorem 3.3 we learn that since $c > 0.5$ the support of the projection $\xi_{\sigma_1^*}$ of the optimal design σ_1^* can be found by calculating the roots of the polynomial $P^*_{d+1}(x)$ defined in (3.5), whereas the weights follow from formula (3.7). Plugging our values in and transforming the result back via (2.6) yields the following symmetric constrained optimal discriminating design

$$\sigma_1^* = \begin{pmatrix} -2.668 & -1.571 & -0.474 & 0.474 & 1.571 & 2.668 \\ 0.1895 & 0.121 & 0.1895 & 0.1895 & 0.121 & 0.1895 \end{pmatrix}.$$

The corresponding value for $\text{eff}_4(\sigma_1^*)$ is given by 0.46, which is the maximal value for this efficiency under the given constraints. The efficiencies for estimating the highest coefficients in the lower order models are obtained as $\text{eff}_1(\sigma_1^*) = 0.4$, and $\text{eff}_3(\sigma_1^*)$ as well as $\text{eff}_2(\sigma_1^*)$ attain the given values for c_3 and c_2. If the constraint $c_3 \geq 0.5$ was relaxed to $c_3 \geq 0.3$ one would obtain the value of 0.66 for the efficiency $\text{eff}_4(\sigma_1^{**})$ when using the corresponding optimal design σ_1^{**}. In this case the lower order efficiencies are given by $\text{eff}_1(\sigma_1^{**}) = 0.4$, and again $\text{eff}_3(\sigma_1^{**})$ and $\text{eff}_2(\sigma_1^{**})$ attain the given values for c_3 and c_2. A higher efficiency eff_4 is thus bought at the expense of a lower efficiency eff_3.
If the other constraint was relaxed to $c_2 = c = 0.4$ we could apply part b) of Theorem 3.3, which yields the optimal design

$$\tilde{\sigma}_1 = \begin{pmatrix} -2.678 & -1.571 & -0.464 & 0.464 & 1.571 & 2.678 \\ 0.156 & 0.188 & 0.156 & 0.156 & 0.188 & 0.156 \end{pmatrix}$$

for this situation. The designs σ_1^* and $\tilde{\sigma}_1$, which are optimal for similar choices of constraints in the same model, have almost the same support points but substantially different weights. It turns out that $\tilde{\sigma}_1$ has efficiencies $\mathrm{eff}_4(\tilde{\sigma}_1^*) = 0.6$, $\mathrm{eff}_3(\tilde{\sigma}_1^*) = 0.4$, $\mathrm{eff}_2(\tilde{\sigma}_1^*) = 0.5$ and $\mathrm{eff}_1(\tilde{\sigma}_1^*) = 0.5$. In this case the lower boundary $c_2 = 0.4$ for $\mathrm{eff}_2(\tilde{\sigma}_1^*)$ is not only attained but even exceeded.

Example 4.2. To illustrate the problem that (symmetric) constrained optimal discriminating designs are not unique we will construct another optimal design (different from σ_1^*) for the situation described in Example 4.1. The projection $\xi_{\sigma_1^*}$ has canonical moments $p_1 = 0.5$, $p_2 = c = 0.6$, $p_3 = 0.5$, $p_4 = 0.23/0.48 \approx 0.479$, $p_5 = 0.5$ and $p_6 = 0$ where the values of $p_1, \ldots, p_4$ follow from Theorem 3.1, and p_5 and p_6 have been added so that this sequence of canonical moments terminates and describes a design on $3 = d+1$ support points within the interval $[-1, 1]$. If one starts with the same sequence $p_1, \ldots, p_4$ of canonical moments but adds different values for the higher order canonical moments, the corresponding design, which is transformed back via (2.6), will be optimal for the same problem as σ_1^*. The efficiencies eff_k, $k = 1, \ldots, 4 = 2d$, will also be the same as for σ_1^* since these values depend on the underlying design only through the canonical moments up to order $4 = 2d$.
Let us consider the sequence of canonical moments $p_1 = 0.5$, $p_2 = c = 0.6$, $p_3 = 0.5$, $p_4 = 0.23/0.48$ and add the values $p_5 = 0.5$ and $p_6 = 1$. From Corollary 4.2.2 in Dette and Studden (1997) it follows that the corresponding design $\xi_{\sigma_1^{new}}$ on the interval $[-1, 1]$ is given by

$$\xi_{\sigma_1^{new}} = \begin{pmatrix} -1 & -0.559 & 0.559 & 1 \\ 0.209 & 0.291 & 0.291 & 0.209 \end{pmatrix}.$$

Transforming this back via (2.6) we obtain the design

$$\sigma_1^{new} = \begin{pmatrix} -\pi & -2.164 & -0.978 & 0 & 0.978 & 2.164 & \pi \\ 0.1045 & 0.1455 & 0.1455 & 0.209 & 0.1455 & 0.1455 & 0.1045 \end{pmatrix},$$

which is optimal for the design problem described in Example 4.1 with the same efficiencies eff_k, $k = 1, \ldots, 4$, as σ_1^*.

Example 4.3. Let us now consider the design problem (2.4) in the situation $d = 3$, $j = 1$ and the pre-specified minimal efficiencies are given by $c_4 = c = 0.6$ and $c_5 = 0.5$. Again, we can apply part a) of Theorem 3.3, which yields the optimal design

$$\sigma_2^* = \begin{pmatrix} -2.793 & -1.906 & -1.235 & -0.349 & 0.349 & 1.235 & 1.906 & 2.793 \\ 0.126 & 0.124 & 0.124 & 0.126 & 0.126 & 0.124 & 0.124 & 0.126 \end{pmatrix}$$

for this scenario. The efficiency for estimating the coefficient of the highest cosine frequency $\mathrm{eff}_6(\sigma_2^*)$, which was to be maximized subject to the given constraints, turns out to be 0.46. Moreover, the efficiencies for estimating the highest coefficients in the lower order models are $\mathrm{eff}_3(\sigma_2^*) = 0.4$, $\mathrm{eff}_2(\sigma_2^*) = \mathrm{eff}_1(\sigma_2^*) = 0.5$, and $\mathrm{eff}_4(\sigma_2^*)$ and $\mathrm{eff}_5(\sigma_2^*)$ attain the given values for c_4 and c_5, respectively.

Example 4.4. Let us finally consider the design problem (2.4) where $d = 3$, $j = 2$, and the constraints are given by $c_4 = c_2 = c = 0.6$ and again $c_5 = 0.5$. So the only difference with respect to Example 4.3 is the additional constraint on the efficiency eff_2. From part a) of Theorem 3.3 we obtain the optimal discriminating design

$$\sigma_3^* = \begin{pmatrix} -2.831 & -1.909 & -1.232 & -0.311 & 0.311 & 1.232 & 1.909 & 2.831 \\ 0.154 & 0.096 & 0.096 & 0.154 & 0.154 & 0.096 & 0.096 & 0.154 \end{pmatrix}.$$

Compared with the optimal design σ_2^* from Example 4.3 the design σ_3^* has similar support points, but the weights differ considerably. The efficiencies are given by $\text{eff}_6(\sigma_3^*) = 0.4$, $\text{eff}_3(\sigma_3^*) = 0.36$, $\text{eff}_1(\sigma_3^*) = 0.4$, and $\text{eff}_2(\sigma_3^*)$, $\text{eff}_4(\sigma_3^*)$ and $\text{eff}_5(\sigma_3^*)$ attain the values of c_2, c_4 and c_5, respectively. The higher value for the efficiency eff_2 is therefore bought at the expense of smaller values for the "most important" efficiency $\text{eff}_6(\sigma_3^*)$ and the efficiencies $\text{eff}_3(\sigma_3^*)$ and $\text{eff}_1(\sigma_3^*)$.

Acknowledgements. This work of the authors was supported by the SFB 475 (Komplexitätsreduktion in multivariaten Datenstrukturen). The authors would also like to thank I. Gottschlich, who typed this paper with considerable technical expertise.

References.

T.W. Anderson (1994). The Statistical Analysis of Time Series. Wiley, Classics Library, NY.

M. Clyde, K. Chaloner (1996). The equivalence of constrained and weighted designs in multiple objective design problems. J. Amer. Statist. Assoc. 91, 1236-1244.

R.D. Cook, W.K. Wong (1994). On the equivalence of constrained and compound optimal designs. J. Amer. Statist. Assoc. 89, 687-692.

H. Dette (1995). Discussion of the paper "Constrained optimization of experimental design" by V. Fedorov and D. Cook. Statistics 26, 153-161.

H. Dette, T. Franke (2000). Constrained D- and D_1-optimal designs for polynomial regression. Annals of Statistics 28, 1702-1727.

H. Dette, G. Haller (1998). Optimal designs for the identification of the order of a Fourier regression. Ann. Statist. 26, 1496-1521.

H. Dette, V.B. Melas (2003). Optimal designs for estimating individual coefficients in Fourier regression models. Annals of Statistics 31, 1669-1692.

H. Dette, V.B. Melas and S. Biedermann (2002). A functional-algebraic determination of D-optimal designs for trigonometric regression models on a partial circle. Stat. Probab. Lett. 58, No.4, 389-397.

H. Dette, V.B. Melas and A. Pepelyshev (2002). D-optimal designs for trigonometric regression models on a partial circle. Ann. Inst. Stat. Math. 54, No.4, 945-959.

H. Dette, W.J. Studden (1997). The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis. Wiley, New York.

V.V. Fedorov (1972). Theory of optimal experiments. Academic Press, New York.

P.D.H. Hill (1978). A note on the equivalence of D-optimal design measures for three rival linear models. Biometrika 65, 666-667.

S. Karlin, W.J. Studden (1966). Tchebycheff systems: with applications in analysis and statistics. Interscience, New York.

C.P. Kitsos, D.M. Titterington, B. Torsney (1988). An optimal design problem in rhythmometry. Biometrics 44, 657-671.

T.S. Lau, W.J. Studden (1985). Optimal designs for trigonometric and polynomial regression. Ann. Statist. 13, 383-394.

C.M.S. Lee (1988a). Constrained optimal designs. J. Statist. Plann. Inf. 18, 377-389.

C.M.S. Lee (1988b). D-optimal designs for polynomial regression, when lower degree parameters are more important. Utilitas Mathematica 34, 53-63.

P.E. Lestrel (1997). Fourier descriptors and their applications in biology. Cambridge Univ. Press.

K. Mardia (1972). The statistics of directional data. Academic Press, New York.

F. Pukelsheim (1993). Optimal Design of Experiments. Wiley, New York.

T.J. Rivlin (1974). Chebyshev polynomials. Wiley, New York.

M. Skibinsky (1967). The range of the $(n+1)$th moment for distributions on $[0,1]$. J. Appl. Prob. 4, 543-552.

M. Skibinsky (1986). Principal representations and canonical moment sequences for distributions on an interval. J. Math. Anal. Appl. 120, 95-120.

S. Stigler (1971). Optimal experimental designs for polynomial regression. J. Amer. Statist. Assoc. 66, 311-318.

W.J. Studden (1968). Optimal designs on Tchebycheff points. Ann. Math. Stat. 39, 1435-1447.

W.J. Studden (1980). D_s-optimal designs for polynomial regression using continued fractions. Ann. Statist. 8, 1132-1141.

W.J. Studden (1982a). Optimal designs for weighted polynomial regression using canonical moments. Third Purdue Symposium on Decision Theory and Related Topics, vol. 2, S.S. Gupta, J.O. Berger eds., 335-350.

W.J. Studden (1982b). Some robust type D-optimal designs in polynomial regression. J. Amer. Statist. Assoc. 8, 916-921.

W.J. Studden (1989). Note on some ϕ_p-optimal designs for polynomial regression. Ann. Statist. 17, 618-623.

M.-M. Zen, M.-H. Tsai (2004). Criterion-robust optimal designs for model discrimination and parameter estimation in Fourier regression models. J. Stat. Plann. Inference 124, 475-487.